28 June 1973

25X1

MEMORANDUM FOR: [redacted]

SUBJECT: Displaying Uncertainty to Decisionmakers

REFERENCE: Memo from A. W. Marshall, 15 June 1973, subject as above, DCI/IC 73-0579

1. Regarding your note attached to Andy Marshall's memo referenced above, this is a matter which has been discussed at one level or another in the intelligence community for more than ten years to my own knowledge, and some experimentation has taken place, but the results are not thusfar impressive. The President's memorandum of November 1971 gave a new impetus with its charge that "a determined effort be made to upgrade analysis personnel and analysis methods."

2. An introduction to quantitative methodologies is provided intelligence personnel in training courses conducted by CIA and DIA, and DIA has an in-house effort going with contractor support, focusing primarily on application of Bayes' theorem.

3. We attempted last year to get a PRG effort going on this problem, to no avail, but perhaps the time is ripe to form a working group and charge it with coming up with some recommendations for application standards and/or a community R&D effort which will result in a set of accepted probability applications and/or methods.

4. What we ought to be looking for is tools which would be useful in the analysis process--recognizing at the same time that obtaining the cooperation and support of analysts and production office managers who are not quantitative-minded is going to be a continuing problem.

5. PRG will work on a specific proposal.

[redacted] 25X1

Deputy Chief, PRG

Attachment
reference

[redacted]: isn't this part of the work [redacted] (?) is undertaking. There is some interesting ARPA work too. Maybe call IDA for ideas — See Rell. [redacted] 7/2 25X1

25X1

25X1

NATIONAL SECURITY COUNCIL
WASHINGTON, D.C. 20506

June 15, 1973

MEMORANDUM FOR

V/ADM. V. P. dePOIX
M/GEN. LEW ALLEN
M/GEN. D. O. GRAHAM
B/GEN. SAMUEL V. WILSON
EDWARD W. PROCTOR
[redacted] 25X1

FROM: A. W. MARSHALL

SUBJECT: Displaying Uncertainty to Decisionmakers

As you know, I have been working with several individuals in efforts to experiment with the use of probabilities and related mathematical tools in community products. Attached is a paper prepared for me by Tom Brown of RAND on one aspect of the subject. You might find it interesting and informative. I remain interested in this method of analysis -- especially in improving the communication of appropriate levels of uncertainty in intelligence judgments. Current methods are unsatisfactory in my view.

DESIRABLE WAYS OF DISPLAYING UNCERTAINTY TO DECISIONMAKERS

T. A. Brown
May 4, 1973

1. Why Should Explicit Probabilities be Used?

Probability is a language which has been worked out over the past three or four hundred years to express varying degrees of uncertainty; it is routinely used in science, insurance, and sports; it is used by systems analysts, hard scientists, actuaries, accountants, demographers, gamblers, soldiers, and a growing body of social scientists; although some of the theory of probability involves difficult abstract mathematics it is not necessary to master the theory in order to use the language meaningfully. One would certainly expect, therefore, that there would be substantial use of probability language in a field as filled with uncertainty as intelligence. By attaching probabilities to alternative ultimately confirmable statements about the world, an intelligence analyst is able to express his inner state of certainty (or uncertainty) with brevity and precision, in a form which can be scored for accuracy (in the long run) and thus contribute to his track record. The purpose of this note is to suggest some patterns, formats, and techniques for expressing intelligence assessments in probabilistic terms which will be easily grouped by the consumer while preserving the inherent potential advantages of brevity, precision, and scorability.

2. What Should Not be Displayed to the Decisionmaker.

Probability theory is a fairly mature branch of mathematical science. It is rich in concepts (such as Bayes' Theorem) and devices (such as event trees and state diagrams) designed to assist in the calculation of probabilities. An assessor is naturally tempted to expose the ingenious and interesting techniques he has employed in deriving his assessments to the appreciative eyes of the decisionmaker. This temptation must be firmly resisted or the brevity of communication will be lost. An event-tree (probability diagram) which fills a page takes longer to absorb than a page of text, and must usually be accompanied by many pages of definition of the states involved and justification for the probabilities ascribed to the various branches. This material is useful to other analysts, perhaps, in discovering how and where differences in

probability assessments arise, but it is no more use to the decisionmaker than would be the worksheets used by the analyst in working out a purely verbal forecast.

The analyst should be careful that the alternatives to which he ascribes probabilities are themselves precisely defined. A probability distribution over vague statements adds up to just another vague statement. If care is not taken in this regard, the statements made to the decisionmaker will be neither precise nor scorable.

A third temptation which must be avoided is the use of excessively "tricky" methods of display. It is often possible, by selecting proper scales or format, to pack a tremendous amount of information into a single picture. What you gain in information density, however, you tend to lose in ease of communication. Therefore, formats should be standardized as much as possible so that the decisionmaker can instantly grasp what is being said. This is especially true for probabilistic statements, where it may be expected that the analyst will be more at home with the language than the decisionmaker.

3. <u>How Probabilities Should be Expressed.</u>

There are an infinite number of mathematically equivalent ways of expressing a probability. We shall discuss six of these, each of which has its own peculiar advantages, and evaluate which of them are most appropriate for probabilistic statements to decisionmakers.

The simplest way to present a probability is to simply state it as a probability. The most obvious advantage of this is simplicity, directness, and the wide understanding of what the word means; a more subtle advantage is that if you are stating the probabilities of more than two disjointed but exhaustive alternatives it is easy to normalize (i. e., the probabilities must add up to 1); the normalization condition on other methods of presentation is much more complex.

Another way of reporting a probability is to give the negative of the logarithm of the probability. This is the amount of "information" in the event. If the logarithm is to the base two, it is expressed in bits. This quantity is called the amount of information in the event because if you have a communication device which has a fixed cost for transmitting a zero or a one (that is, a fixed cost per bit), then you will minimize your expected cost of communication by selecting a code which assigns to each

Name	Symbol	Range	Formula	Advantages
Probability	p	$0 \le p \le 1$	p	Common use; easy normalization
Log Probability	I	$0 \le I \le \infty$	$I = -\log_2 p$	Information theory; scoring system
Odds	d	$0 \le d \le \infty$	$d = \frac{p}{1-p}$	Common use; wagers
Log Odds	L	$-\infty \le L \le \infty$	$L - \log d$	Bayes' theorem
Pari-mutual Payoff	M	$1 \le M \le \infty$	$M = \frac{1}{p}$	Use at race tracks
Insurance Premium	P	$0 \le p \le 1$	$P = p$	A probability is a price

Six Equivalent Ways of Expressing a Probability

TABLE 1

Probability	Information	Odds	Log Odds	Parimutual Payoff	Premium
0	$+\infty$	0	$-\infty$	$+\infty$	0
.05	4.32	.053	-1.28	20.	.05
.1	3.32	.111	- .95	10.	.1
.2	2.32	.25	- .60	5.	.2
.3	1.74	.43	- .37	3.33	.3
.4	1.32	.67	- .18	2.5	.4
.5	1.00	1.00	0	2.	.5
.6	.74	1.5	.18	1.67	.6
.7	.51	2.33	.37	1.43	.7
.8	.32	4.	.60	1.25	.8
.9	.15	9.	.95	1.11	.9
.95	.07	19.	1.28	1.05	.95
1.	0	$+\infty$	$+\infty$	1.	1.

A Short Table of Values of Alternative Ways of Expressing a Probability

TABLE 2

of the possible events you may have to report a signal containing a number of bits equal to the negative $\log_2$ of the probability of the event. In other words, if you have selected an optimal code, $-\log_2 p$ is the number of bits you will have to send to communicate the occurrence of an event whose probability is p. The reader familiar with reproducing scoring systems will recognize the logarithm of the probability as a scoring system which encourages a respondent to be "honest" (i. e., to report his true subjective probability that a given event will take place). These characteristics make the log of the probability a useful training device for analysts and of great interest to information theorists, but I suspect it would tend to confuse a decisionmaker who is not very familar with these disciplines.

A more down-to-earth way of expressing a probability is to translate it into the equivalent betting odds. The higher the odds, the more likely the event is to occur. If the odds on an event are 2:1, then any chance to bet on the event at odds of 1.9:1 should be accepted. A bet at odds at 2.1:1 should be rejected. Some individuals are very comfortable thinking in terms of odds. I understand from one person with a great deal of experience teaching probability theory to analysts that many individuals find odds repugnant (perhaps because of their association with frivolous games of chance) while they find probabilities (which have a much more "scientific" aura) perfectly acceptable.

Some Bayesians (notably Ward Edwards) have advocated encoding probabilities by means of the logarithm of the odds. The advantage of this is that under Bayes' Theorem the change in the log odds ratio if a new piece of evidence comes in depends only on the new evidence, not on the prior value of the log-odds ratio. To be perfectly explicit, let us suppose that we have a prior probability p that hypothesis H is true (and thus probability 1-p that H is not true). Suppose an event E takes place, with:

$p(E/H)$ = probability that E occurs if H is true.

$p(E/\sim H)$ = probability that E occurs if H is not true,

then by Bayes' Theorem the posterior odds will be:

$$\frac{p}{1-p} \times \frac{p(E/H)}{p(E/\sim H)}$$

Thus the log odds ratio will change by an amount

$$\log \frac{p(E/H)}{p(E/\sim H)}$$

regardless of what its initial value may have been. If you have a situation in which probabilities are continually being adjusted by applying Bayes' Theorem with new information flowing in, then it is easy to see that there are advantages to encoding your probabilities in this way. On the other hand, although the decisionmaker may be interested in getting information in probabilistic terms, he is not apt to be called upon to perform computational manipulations on them. Therefore I think that this particular method of encoding is not appropriate for communicating with the decisionmaker.

A fifth way of encoding probabilities is to calculate the amount you would be paid if you should win a $1.00 for a bet placed with a bookie on the event. This is simply the reciprocal of the probability of the event. For individuals whose major contact with decisionmaking under explicit uncertainty takes place at race tracks, this quantity (perhaps multiplied by two) will provide quite a meaningful expression for the likelihood of a given event. If the decisionmaker in question falls into this category, then the use of this encoding may have something to recommend it.

A final way of encoding a probability is to consider it a price: the premium you would have to pay for a one-dollar insurance policy against the event taking place. This price is numerically equal to the probability of the event, so you might argue that this encoding is simply identical to the first one named above: the probability itself. But psychologically it is a little different. Viewing a probability as a price (an idea due, I believe, to Savage) makes it obvious that it makes sense to speak of the probability of inherently unique events. Many people, who have been exposed to a bit of positivist epistemology, hold that to speak of a probability you must either have some sub-set of a set of equally likely cases (combinational probability) or else an inherently repeatable case (such as flipping a coin, for example). Calling a probability a price helps this sort of person overcome their inhibitions against applying the concepts in the context of elections, resolutions, wars, economic change, and all the other unique events which characterize the world of the national security decisionmaker.

To summarize the six encoding schemes above, let us consider identically the same piece of information encoded in six different ways:

"The probability that Broyhill will be re-elected is .7."

> "The information content of Broyhill's re-election would be about half a bit."
>
> "I will lay \$2.33 to \$1.00 that Broyhill is re-elected."
>
> "The log odds (base 10) on Broyhill's re-election are +.37."
>
> "A two dollar bet on Broyhill's re-election pays \$2.86."
>
> "A \$1,000 insurance policy against Broyhill's re-election costs \$700."

Which of these is to be preferred? I like the first and last best: they are the most straight-forward. Next I like odds, and after that the betting house pay-off. The encodings involving logarithms are too esoteric for the decisionmaker unless he intends to do some computation, which seems highly unlikely to me.

4. <u>"Futures" Charts</u>.

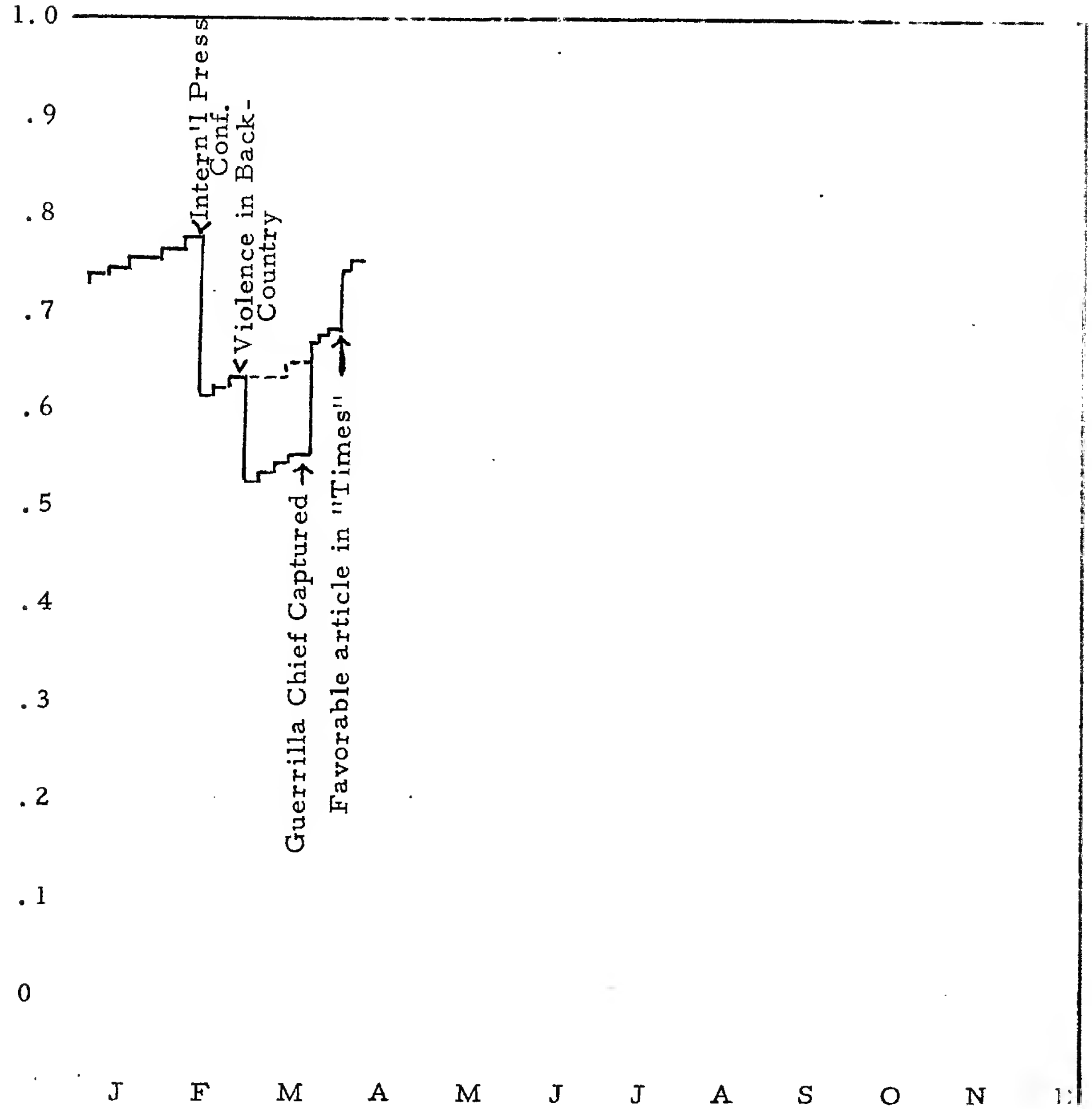
The futures market in any commodity is a predictive device. It is a forecast of what the price of a given commodity is going to be at some specific date in the future. Anyone interested in the commodity will study the futures market and the movement of future prices with great care, for it indicates a consensus of views on the probable supply and demand over the immediate future.

Couldn't we set up a similar mechanism for forecasting events other than the movement of commodity prices? For example, a market in insurance policies that General Flup will (or will not) be premier of Ruritania on December 31, 1973? If we would try to set up an actual market, the administrative burden of keeping track on who sold what and who bought what would outweigh the benefits of the operation, but we could set up an excellent simulacrum of such a market by exploiting reproducing scoring systems. For example, if we had three experts on Ruritania we could ask each of them to produce by each Monday noon an estimate of the probability the General Flup will be premier of Ruritania on December 31, 1973, together with a brief explanation of what factors in the situation have caused them to change their estimate from the preceding Monday. At the end of the year each of the three experts would be individually rewarded proportionally to the sum of the scores of their weekly forecasts (note: it is important <u>not</u> to give any "special"

reward to the forecaster achieving the highest score, as this would disrupt the incentive system and tend to induce excessively "risky" estimates). If a crisis flares up, the experts could, of course, begin revising their estimates daily or even hourly.

Making such estimates on a systematic schedule, and graphing them up appropriately (see Figure 3) has many advantages. First, when such a chart has been established and the decisionmaker has taken the trouble to understand it once, he will be able to absorb its message very rapidly whenever he looks at it in the future. Second, the chart automatically generates a chronology of relevant events and indications and gives a subjective grading of their relative importance (an important event corresponds to a big jump in the graph: an unimportant one to a small jump). Third, graphs such as these are very "scorable" (generating, as they do, fifty or more responses from each analyst on each question). Fourth, they provide a ready vehicle for expressing "minority opinion" by analysts. For example, in the chart shown (Figure 3) one analyst did not drop the probability of General Flup remaining in office in response to reports of back-country violence, while the others did. The minority analysts' assessment is plotted separately from the mean of the majority for a time. This is quite an advantage of quantitative assessments of uncertainty: where a disagreement arises, it is very easy to report in just how much it amounts to and report it precisely. There is no need to waste time trying to create some vague verbal formula which covers all points of view or to draft reclamas. Furthermore, since each analyst is held responsible for his own assessments at scoring time there is motivation to dissent when and only when you feel you are right and the majority is wrong (if the assessments are prepared independently, as they ought to be, the "dissent" is probably not the right word; the important point is that since the analysts are each individually responsible for their individual assessments, rather than being held collectively responsible for a group assessment, individual deviations will occur fairly frequently and will be ultimately punished or rewarded as they are less or more accurate than the group as a whole).

Another example of such a graph may be found on page 23 of Peterson, Kelly, Barclay, Hazard Handbook for Decision Analysis: Inference from Evidence: Bayes' Theorem. This example is presented to illustrate the application of Bayes' Theorem, but such a display is useful as a communication device whether you use Bayes or some other statistical technique or just informed intuition as the basis for making your forecasts. My personal prejudice is that the informed intuition of



Graph of "Event Futures" - Probability General Glup will will be Premier of Ruritania on December 31, 1973

FIGURE 3

*One analyst feels that the back-country violence is not serious, and indeed has no political significance. His estimate is indicated by the dashed line.

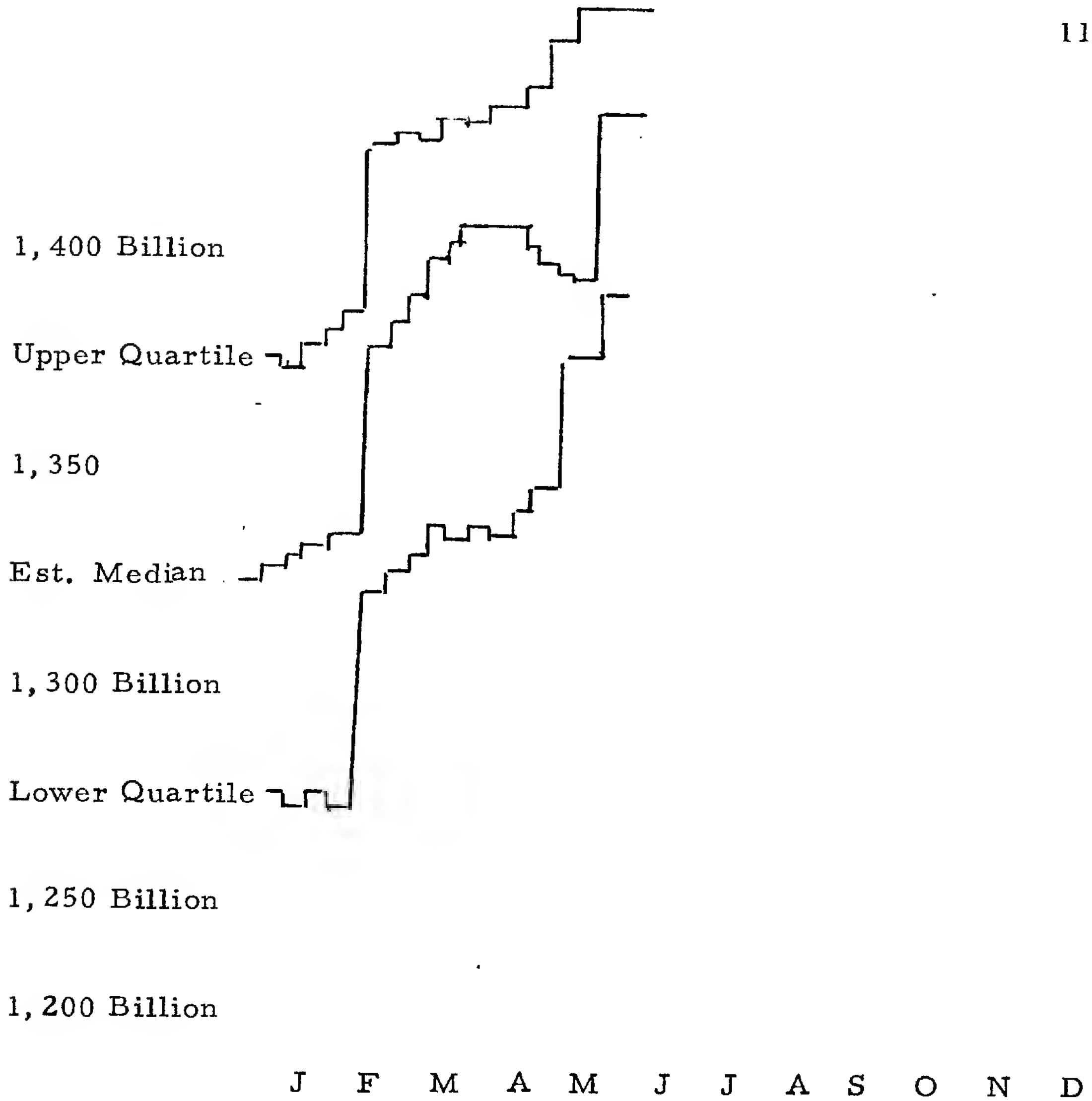
a person properly trained in using the language of probabilities will usually turn out to be the most sensitive and accurate instrument available in arriving at probability estimates of real-world events; but this is an empirical matter, and it may turn out in real cases that people who use formal methods consistently out perform those who do not.

Some questions on which you may want to keep futures charts are difficult to express as probabilities of single events. For example, let us suppose you wish to forecast the GNP of the United States for 1973. You could cast this as "What is the probability that the GNP will exceed 1.3 trillion dollars in 1973?" but this format would clearly lack resolution. A better format would be to ask each analyst to estimate (say) the median, upper quartile, lower quartile estimate of what the GNP will turn out to be. If he makes such an estimate each week, then you can grade him at the end of the year by using some recipe to construct a distribution from the three numbers he gave you and applying one at the reproducing scoring systems appropriate to continuous distributions. The distributions generated by different analysts may be combined in any of a number of ways, and as in the case of single event probabilities any distribution which varies too much from the "combined" distribution may be displayed separately. The graph would look something like Figure 4.

Some will complain that keeping charts of this kind is apt to turn out, in many cases, to be rather a waste of time because something which looks like a hot issue or big problem when the chart is set up may recede from interest before the chart reaches completion. Well, if this happens the completion of the chart will be a very mild burden on the analyst (if General Flup dies, for instance, the analyst hardly has to pore over the dispatches from Ruritania to decide what his chances at the premiership might be). In any case, a single analyst could be responsible for providing inputs to 10 or 12 charts each week. If an average of three analysts are feeding data to each chart, then this means that we will be able to keep 3.5 to 4 charts going for each analyst assigned to this duty. Thus 50 analysts ought to provide enough brain-power to keep 175 - 200 charts going.

5. Point Estimates.

Almost any intelligence report which contains an element of prognostication can be recast in explicit probabilistic terms. If you take all the implicit probabilistic forecasts in a paper, make them explicit, and simply list



Consensus of Probabilistic Estimates of U.S. GNP
for 1973

FIGURE 4

them, this will often turn out to constitute an excellent executive summary of the forecasting element of the paper. Having prepared such a summary an easy way to ask some group other than the authors of the paper for comments is to simply ask them what probabilities they would ascribe to the various events in question.

If you do this in a systematic manner, you will soon (in a year or so) begin to get some objective measure of the relative accuracy of CIA, DIA, and any other intelligence service you have available. I think it is very dangerous for a decisionmaker to evaluate his intelligence services on the grounds of whether they structure the problem the way he does or not. If he is right and the intelligence service is wrong a couple of times he may easily get carried away by hubris and commit some really big errors. Intelligence should be rated for accuracy and relevance, not for how well it fits the decisionmaker's view of the world.

25X1